PCT/EP 0 3 / 1 2 2 6 1

Europäisches Patentamt | European Patent Office | Office européen des brevets

Bescheinigung Certificate Attestation

Die angehefteten Unterlagen stimmen mit der ursprünglich eingereichten Fassung der auf dem nächsten Blatt bezeichneten europäischen Patentanmeldung überein.

The attached documents are exact copies of the European patent application described on the following page, as originally filed.

Les documents fixés à cette attestation sont conformes à la version initialement déposée de la demande de brevet européen spécifiée à la page suivante.

Patentanmeldung Nr. Patent application No. Demande de brevet n°

02025474.4

Der Präsident des Europäischen Patentamts;
Im Auftrag

For the President of the European Patent Office

Le Président de l'Office européen des brevets
p.o.

R C van Dijk

Anmeldung Nr:
Application no.: 02025474.4
Demande no:

Anmeldetag:
Date of filing: 15.11.02
Date de dépôt:

Anmelder/Applicant(s)/Demandeur(s):

DEUTSCHE THOMSON-BRANDT GMBH
Hermann-Schwer-Strasse 3
78048 Villingen-Schwenningen
ALLEMAGNE

Bezeichnung der Erfindung/Title of the invention/Titre de l'invention:
(Falls die Bezeichnung der Erfindung nicht angegeben ist, siehe Beschreibung.
If no title is shown please refer to the description.
Si aucun titre n'est indiqué se referer à la description.)

Method and apparatus for improved pixel-based subtitling

In Anspruch genommene Priorität(en) / Priority(ies) claimed /Priorité(s) revendiquée(s)
Staat/Tag/Aktenzeichen/State/Date/File no./Pays/Date/Numéro de dépôt:

Internationale Patentklassifikation/International Patent Classification/
Classification internationale des brevets:

H04N/

Am Anmeldetag benannte Vertragstaaten/Contracting states designated at date of filing/Etats contractants désignées lors du dépôt:

AT BE BG CH CY CZ DE DK EE ES FI FR GB GR IE IT LI LU MC NL PT SE SK TR

Method and Apparatus for improved pixel-based subtitling

The invention relates to a method and to an apparatus for improved pixel-based subtitling, which can be used e.g. for HDTV subtitles in pre-recorded formats like the so-called Blue-ray Disc.

Background

The technique of subtitling for Audio-Visual material has been used beginning with the first celluloid cinema movies and further until the recent digital media appeared. The main target of subtitling has been the support of handicapped people or small ethnographic language groups. Therefore subtitling often aims at the presentation of text information even when having been encoded as graphic data like pixel maps. Therefore pre-produced AV-material for broadcasting (Closed Caption, Teletext, DVB-Subtitle etc.) and movie discs (DVD Sub-Picture etc.) primarily are optimized for subtitles representing simple static textual information. However, progress in PC software development for presentation and animation of textual information induces a corresponding demand for possibilities and features within the digital subtitling technique used for pre-recording and broadcasting. Using straightforward approaches without any special precautions, these increased requirements for subtitling would consume a too big portion of the limited overall bandwidth. The conflicting requirements for a 'full feature' subtitle encompassing karaoke all through genuine animations are on one hand the coding efficiency and on the other hand the full control for any subtitle author.

For today's state of the art of digitally subtitling AV material with separate subtitling information two main approaches exist : Subtitling can be based on either pixel

data or on character data. In both cases, subtitling schemes comprise a general framework, which for instance deals with the synchronization of subtitling elements along the AV time axis.

Character data based subtitling:
In the character-based subtitling approach, e.g. in the teletext system [1] of European analog or digital TV, strings are described by sequences of letter codes (e.g. ASCII [2] or UNICODE [3]), which intrinsically allows for a very efficient encoding. But from character strings alone, subtitling can not be converted into a graphical representation to be overlaid over video. For this, the intended character set, font and some font parameters, most notably the font size, must either be coded explicitly within the subtitling bitstream or an implicit assumption must be made about them within a suitably defined subtitling context. Also, any subtitling in this approach is confined to what can be expressed with the letters and symbols of the specific font(s) in use. The DVB Subtitling specification [4], in its mode of "character objects", constitutes another state-of-the-art example of character-based subtitling.

Pixel data based subtitling:
In the pixel-based subtitling approach, subtitling frames are conveyed directly in the form of graphical representations by describing them as (typically rectangular) regions of pixel values on the AV screen. Whenever anything is meant to be visible in the subtitling plane superimposed onto video, its pixel values must be encoded and provided in the subtitling bitstream, together with appropriate synchronization info and hence for the full feature animation of subtitles all pixel changed must be transported. Obviously when removing any limitations inherent with full feature animations of teletext, the pixel-based approach carries the penalty of a considerably increased bandwidth for the subtitling data. Examples of pixel-based subtitling schemes can

be found in DVD's "Sub-picture" concept [5] as well as in the "pixel object" concept of DVB Subtitling [4].

Invention

The gist of the invention is a subtitling format encompassing elements of enhanced syntax and semantic to provide improved animation capabilities. The disclosed elements improve subtitle performance without stressing the available subtitle bitrate. This will become essential for authoring content of high-end HDTV subtitles in pre-recorded format which can be broadcast or pressed on high capacity optical media, e.g. the Blue-ray Disc. The invention includes abilities for improve authoring possibilities for the content production to animate subtitles.

Introduced by the disclosure are elements of syntax and semantic describing the color change for parts of graphics to display. This can be used for highlight effects in applications like for example karaoke, avoiding the repeated transfer of pixel data.

Other disclosed elements of syntax and semantic facilitate the ability of cropping parts of the subtitles before displaying them. By using the technique of subsequent transferred cropping parameters for an object to display a bit saving animation of subtitles becomes available. Such cropping parameter can be used for example to generate text changes by wiping boxes, blinds, scrolling, wipes, checker boxes, etc.

Furthermore the disclosed elements can be used to provide interactivity on textual and graphical information. Especially the positioning and/or color settings of subtitles can be manipulated based upon user request.

Drawings

Exemplary embodiments of the invention are described with reference to the accompanying drawings and tables, which show:

Figure 1: Example for the definition of a subtitle region and its location within a page;
Figure 2: Example for definition of a region sub-CLUT and region cropping;
Figure 3: Resulting display example;
Figure 4: Interactive usage of subtitles;
Table 1: segment_type values for enhanced PCS and RCS;
Table 2: Enhanced page composition segment;
Table 3: Enhanced region composition segment;
Figure 5: Video and Graphics Planes;
Figure 6: Video and Graphics Mixing and Switching.

Exemplary embodiments

The invention can preferably be embodied based on the syntax and semantic of the DVB subtitle specification (DVB-ST) [4]. To provide improved capabilities for the manipulation of graphic subtitle elements the semantics of DVB-ST's page composition segment (PCS) and region composition segment(RCS) are expanded.
DVB_ST uses page composition segments (PCS) to describe the positions of one or more rectangular regions on the display screen. The region composition segments (RCS) are used to define the size of any such rectangular area and identifies the color-lookup-table used within.
The proposed invention keeps backward compatibility with DVB-ST by using different segment_types for the enhanced PCS and RCS elements as listed in Table 1. It would of course be possible to choose other values instead. Another approach for keeping backward compatibility would be to keep the existing segment_types and increase the version_number of the specification (e.g. by incrementing the subti-

tle_stream_id in the PES_data_field structure). The enhanced PCS is defined as listed in Table 2 and the enhanced RCS is defined as listed in Table 3. With respect to original DVB-ST, all structures shown are expanded, in the tables the additional entries are marked gray.

The enhanced PCS carries optional information about the region cropping and optional information about the region sub-CLUT for every region listed. The 2 values of region_cropping and region_sub_CLUT indicate if such optional information is available for the current region in process. Therefore cropping and sub-CLUT may be defined separately for every region. While region_cropping is used as a flag (if equal 0x01), the region_sub_CLUT shows the value how many sub-CLUT positions are described. This is done to provide different alternatives within the stream. Alternative sub-CLUT positions can be used to define different menu button positions for the display screen. Only one of them - the first one as a default - is active and the user can change the position to navigate through the different predefined positions pressing the remote for example.

The enhanced RCS carries the sub_CLUT_id identifying the family of CLUTs that applies to this region. This is done to re-use CLUTs for different regions and different region sub_CLUTs as well.

The enhanced PCS and enhanced RCS elements provide the ability that subtitles can be manipulated independent from the encoding method i.e. independent from whether they are encoded as character data or pixel data.

The enhanced PCS and RCS can be used to perform many different animation effects for subtitles. Those could be wiping boxes, blinds, scrolling, wipes, checker boxes, etc. The following Figures show an application example for karaoke. Figure 1 shows the definition of a region containing lyrics of a song displayed for karaoke. The letters of the subtitle may be encoded as pixel data or as character data as well. The region_vertical_address and region_horizontal_address define the location of the subtitle

within the frame to display.

Figure 2 depicts the region cropping and the location of the region sub-CLUT.

The region cropping defines which part of the region is effectively displayed. This is achieved by 4 parameters indicating the start coordinates and the size of the fragment to display. region_horizontal_cropping specifies the horizontal address of the top left pixel of this cropping and region_vertical_cropping specifies the vertical address of the top line of this cropping. region_cropping_width specifies the horizontal length of this cropping and region_cropping_height specifies the vertical length of this cropping.

The region sub-CLUT location defines which part of the region has to be displayed using a color look up table different from the region CLUT. This is achieved by 4 parameters indicating the start coordinates and the size of the sub-region used by the sub-CLUT. All coordinate parameters are to be understood relative to the region the sub-CLUT belongs to. sub_CLUT_horizontal_address specifies the horizontal address of the top left pixel of this sub-CLUT and sub_CLUT_vertical_address specifies the vertical address of the top line of this sub-CLUT. sub_CLUT_width specifies the horizontal length of this sub-CLUT and sub_CLUT_height specifies the vertical length of this sub-CLUT.

Picking up all parameters defined with the previous Figures results in the displayed subtitle as depicted in Figure 3. The subtitle is not depicted in whole but only a part of it. Furthermore the sub-CLUT was used to provide a highlight, so that the user knows what to sing in the moment.

As the enhanced PCS are sent within MPEG PES packets labeled by presentation time stamps PTS any effect can be synchronized to the AV.

Another idea of the invention is the superseding of subtitle animation parameters by the user. This offers a way to realize interactive subtitles. The enhanced PCS parameters are transferred as a default and the user may change them

via a remote control for example. By this the user is able to move, crop or highlight the subtitle.

This could be an advantage for a user defined repositioning of a subtitling text, so that the user can subjectively minimize the annoyance by the subtitle text placement on top of the motion video. Also the color of the subtitles could be set according to users preferences. Figure 4 shows a block diagram for interactive subtitle modifications. The default parameters are superseded by the user action.

Another application for overriding subtitle animation parameters like position, cropping rectangle, cluts, sub-cluts is the realization of some very basic sort of interactive gaming. The subtitle may carry pixel data of an animated character. This character is subsequently moved on the display screen driven by either user interaction, programmatic control or both.

The overriding of subtitle animation parameters can be implemented in at least two ways. The first option is that the overriding parameters replace the parameters send in the bitstream. The second option is that the overriding parameters are used as an offset that is added to or subtracted from the subtitle animation parameters send in the bitstream.

The enhances PCS and RCS provide a lot more of animation capabilities not explained. Following is a non-exhaustive list of exemples: wiping boxes, blinds, scrolling, wipes, checker boxes in details.

References:

[1] Teletext – ETS 300 706: Enhanced Teletext Specification

[2] ASCII - ISO/IEC 8859: American Standard Code for Information Interchange - ASCII

[3] UNICODE – ISO/IEC 10646: Information technology -- Universal Multiple-Octet Coded Character Set (UCS)

[4] DVB Subtitling - ETS 300 743: Digital Video Broadcasting (DVB); Subtitling System

[5] DVD Sub-picture - DVD Specification for Read-Only Disc: Part 3 - Video Specification

[6] PNG - PNG (Portable Network Graphics) specification; a W3C Recommendation on 1st October, 1996

[7] MNG - MNG (Multiple-image Network Graphics) specification; approved by the W3C PNG specification Group

Claims

1. Method for improved pixel-based subtitling, wherein subtitling frames are conveyed in the form of graphical representations by describing them as regions of pixel values, wherein optional information about cropping and/or color change for parts of a region is provided.

2. Apparatus for performing a method according to claim 1.

Abstract

The gist of the invention is a subtitling format encompassing elements of enhanced syntax and semantic to provide improved animation capabilities. The disclosed elements improve subtitle performance without stressing the available subtitle bitrate. This will become essential for authoring content of high-end HDTV subtitles in pre-recorded format which can be broadcast or pressed on high capacity optical media, e.g. the Blue-ray Disc. The invention includes abilities for improve authoring possibilities for the content production to animate subtitles.

Fig. 2

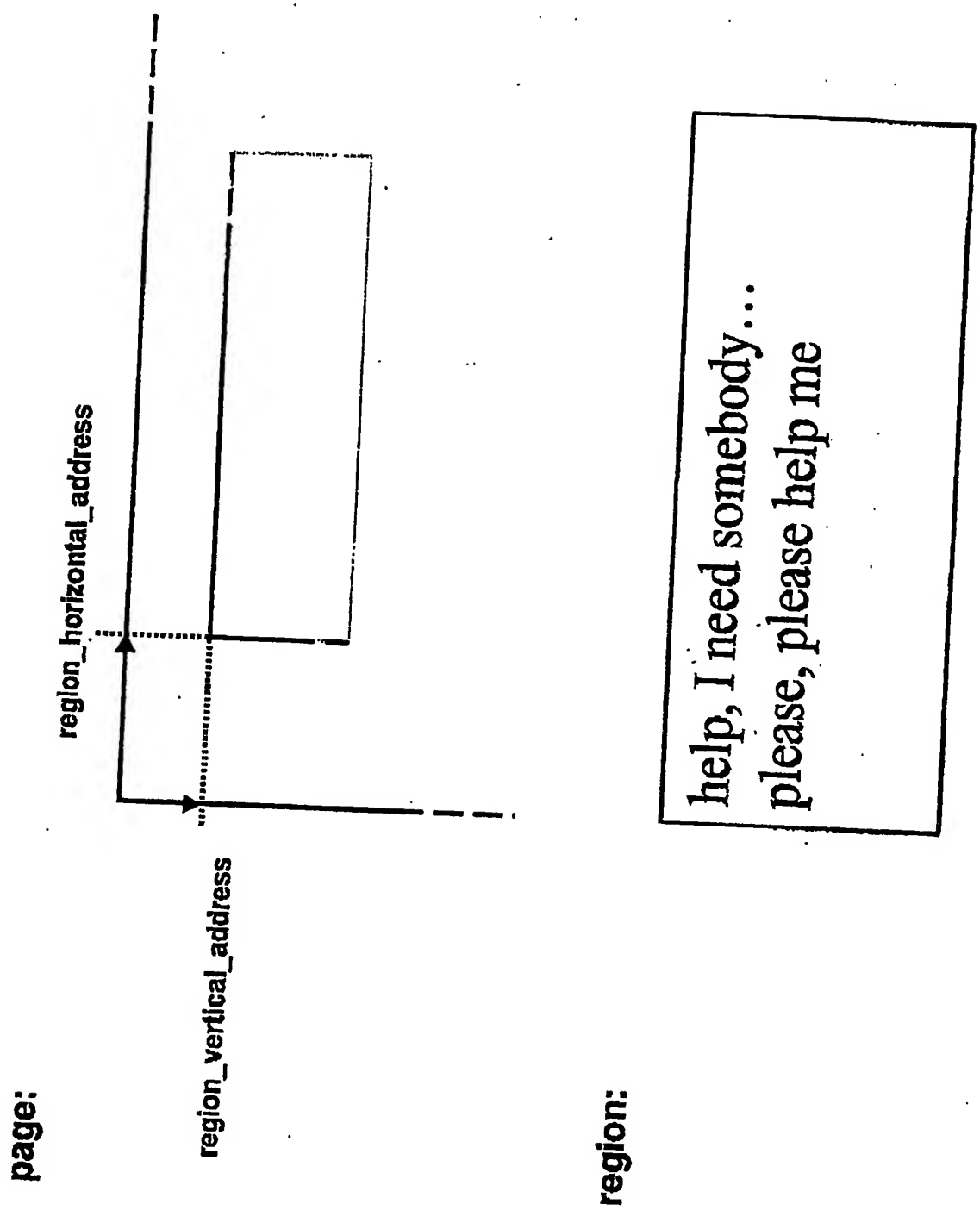

region_horizontal_address
region_vertical_address
page:
region:
help, I need somebody...
please, please help me

Figure 1

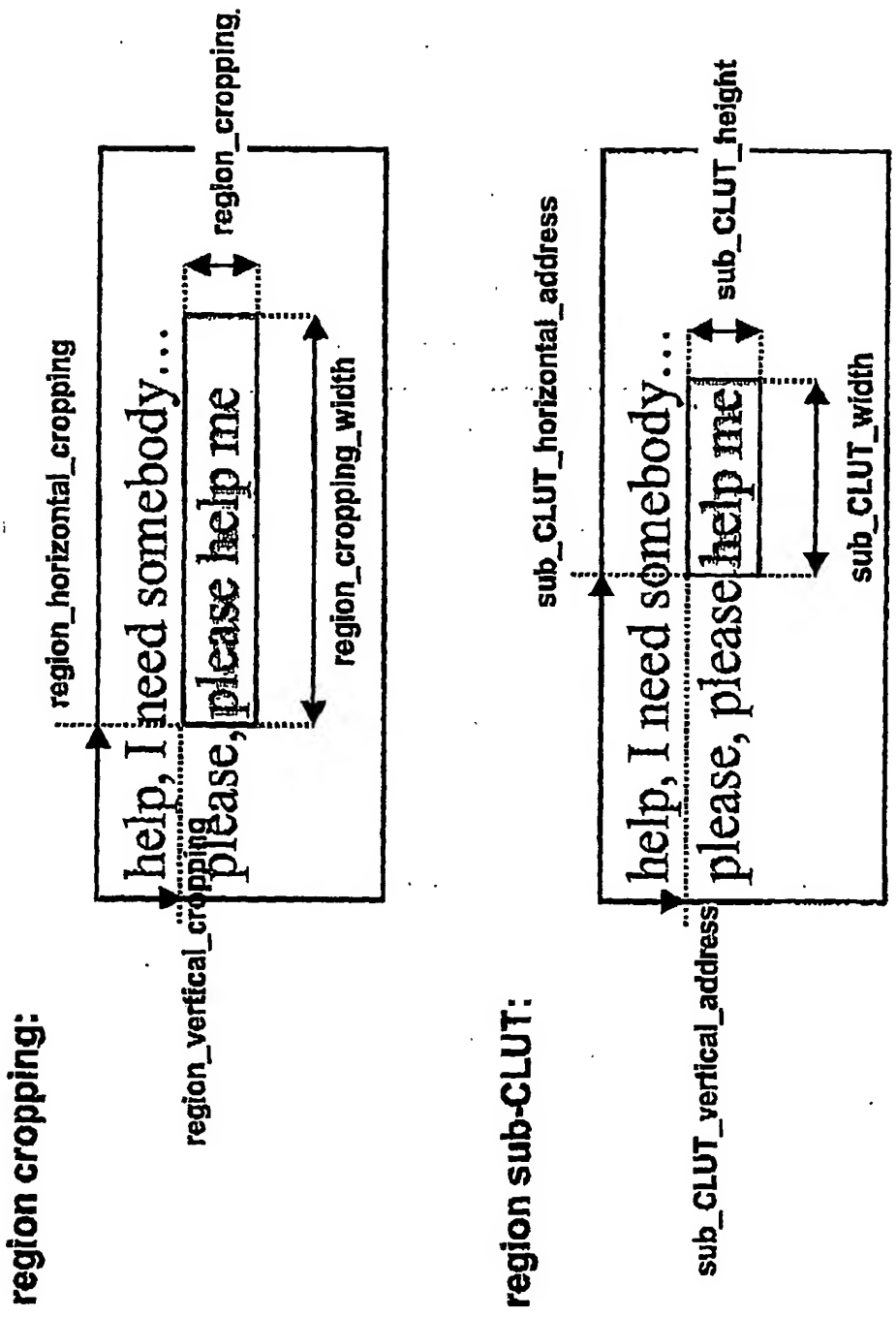
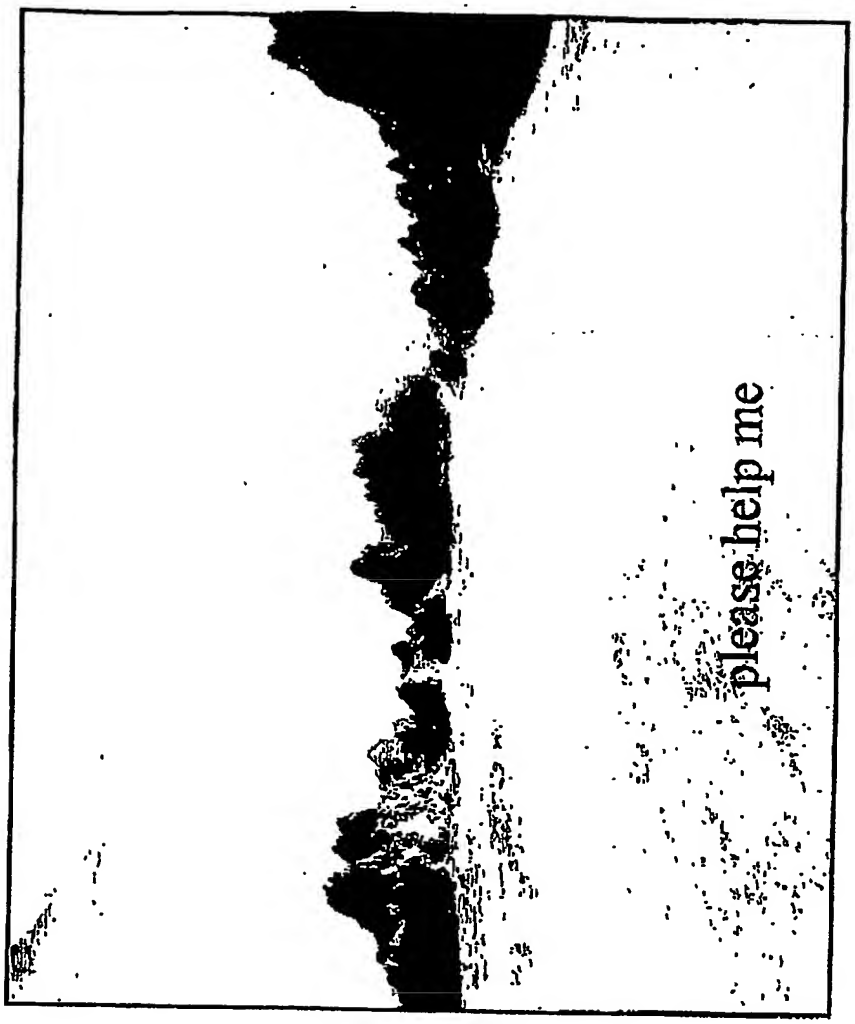
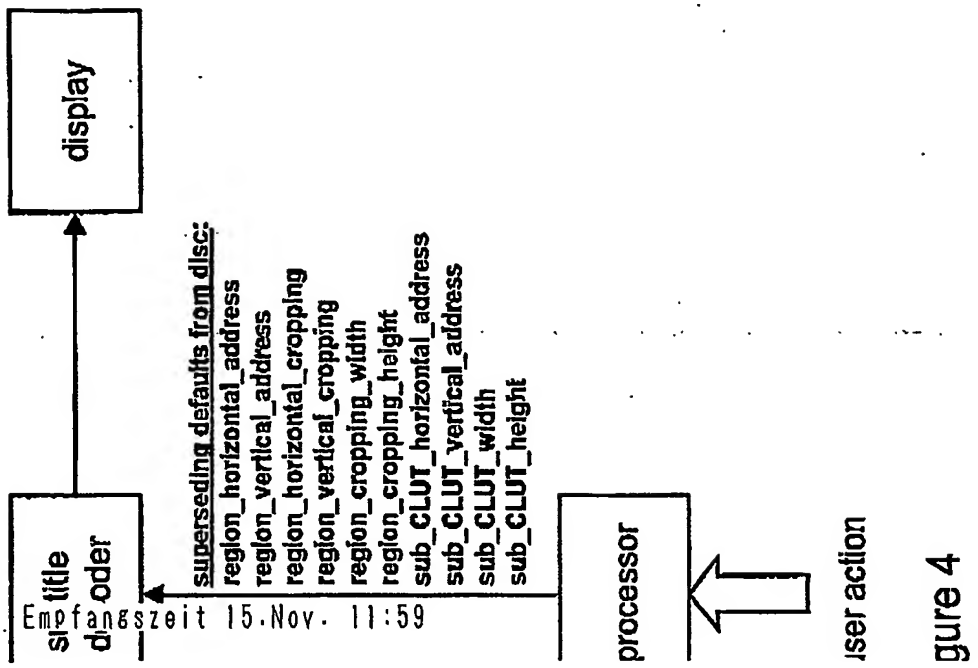

region cropping:
region_horizontal_cropping
region_cropping
region_vertical_cropping
region_cropping_width
help, I need somebody...
please, please help me
region sub-CLUT:
sub_CLUT_horizontal_address
sub_CLUT_height
sub_CLUT_vertical_address
sub_CLUT_width
help, I need somebody...
please, please help me

Figure 2



display:

Figure 3


display
superseding defaults from disc:
region_horizontal_address
region_vertical_address
region_horizontal_cropping
region_vertical_cropping
region_cropping_width
region_cropping_height
sub_CLUT_horizontal_address
sub_CLUT_vertical_address
sub_CLUT_width
sub_CLUT_height
Empfangszeit 15.Nov. 11:59
processor
ser action

gure 4

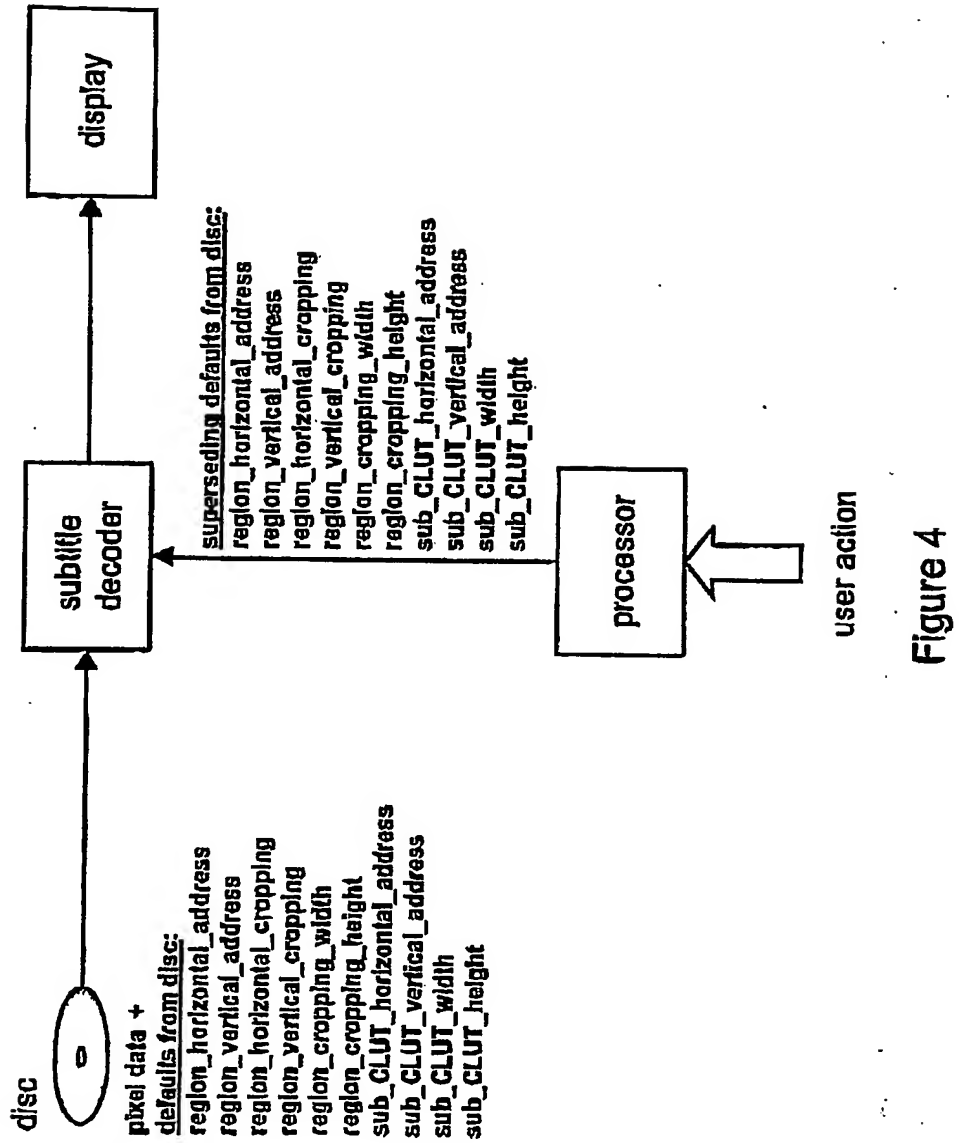

display
subtitle decoder
processor
user action
disc
pixel data +
defaults from disc:
region_horizontal_address
region_vertical_address
region_horizontal_cropping
region_vertical_cropping
region_cropping_width
region_cropping_height
sub_CLUT_horizontal_address
sub_CLUT_vertical_address
sub_CLUT_width
sub_CLUT_height
superseding defaults from disc:
region_horizontal_address
region_vertical_address
region_horizontal_cropping
region_vertical_cropping
region_cropping_width
region_cropping_height
sub_CLUT_horizontal_address
sub_CLUT_vertical_address
sub_CLUT_width
sub_CLUT_height

Figure 4

0x10	page composition segment	defined in 7.2.1.
0x11	region composition segment	defined in 7.2.2.
0x12	CLUT definition segment	defined in 7.2.3.
0x13	object data segment	defined in 7.2.4.
0x14	enhanced page composition segment	defined here
0x15	enhanced region composition segment	defined here
0x40 - 0x7F	reserved for future use	
0x80	end of display set segment	defined in 7.2.5.
0x81 - 0xEF	private data	
0xFF	stuffing	
All other values	reserved for future use	

Table 1

Syntax	Size	Type
page_composition_segment() {		
sync_byte	8	bslbf
segment_type	8	bslbf
page_id	16	bslbf
segment_length	16	uimsbf
page_time_out	8	uimsbf
page_version_number	4	uimsbf
page_state	2	bslbf
reserved	2	bslbf
while (processed_length < segment_length) {		
region_id	8	bslbf
reserved	8	bslbf
region_horizontal_address	16	uimsbf
region_vertical_address	16	uimsbf
region_cropping	8	bslbf
if (region_cropping == '0x01') {		
region_horizontal_cropping	16	uimsbf
region_vertical_cropping	16	uimsbf
region_cropping_width	16	uimsbf
region_cropping_height	16	uimsbf
}		
region_sub_CLUT	8	uimsbf
for (i=0; i<region_sub_CLUT; i++) {		
sub_CLUT_horizontal_address	16	uimsbf
sub_CLUT_vertical_address	16	uimsbf
sub_CLUT_width	16	uimsbf
sub_CLUT_height	16	uimsbf
}		
}		
}		

Table 2

Syntax	Size	Type
region_composition_segment () {		
sync_byte	8	bslbf
segment_type	8	bslbf
page_id	16	bslbf
segment_length	16	uimsbf
region_id	8	uimsbf
region_version_number	4	uimsbf
region_fill_flag	1	bslbf
reserved	3	bslbf
region_width	16	uimsbf
region_height	16	uimsbf
region_level_of_compatibility	3	bslbf
region_depth	3	bslbf
reserved	2	bslbf
CLUT_id	8	bslbf
sub_CLUT_id	8	bslbf
region_8-bit_pixel-code	8	bslbf
region_4-bit_pixel-code	4	bslbf
region_2-bit_pixel-code	2	bslbf
reserved	2	bslbf
while (processed_length < segment_length) {		
region_id	8	bslbf
[...]		
}		
}		

Table 3

Video and Graphics Planes

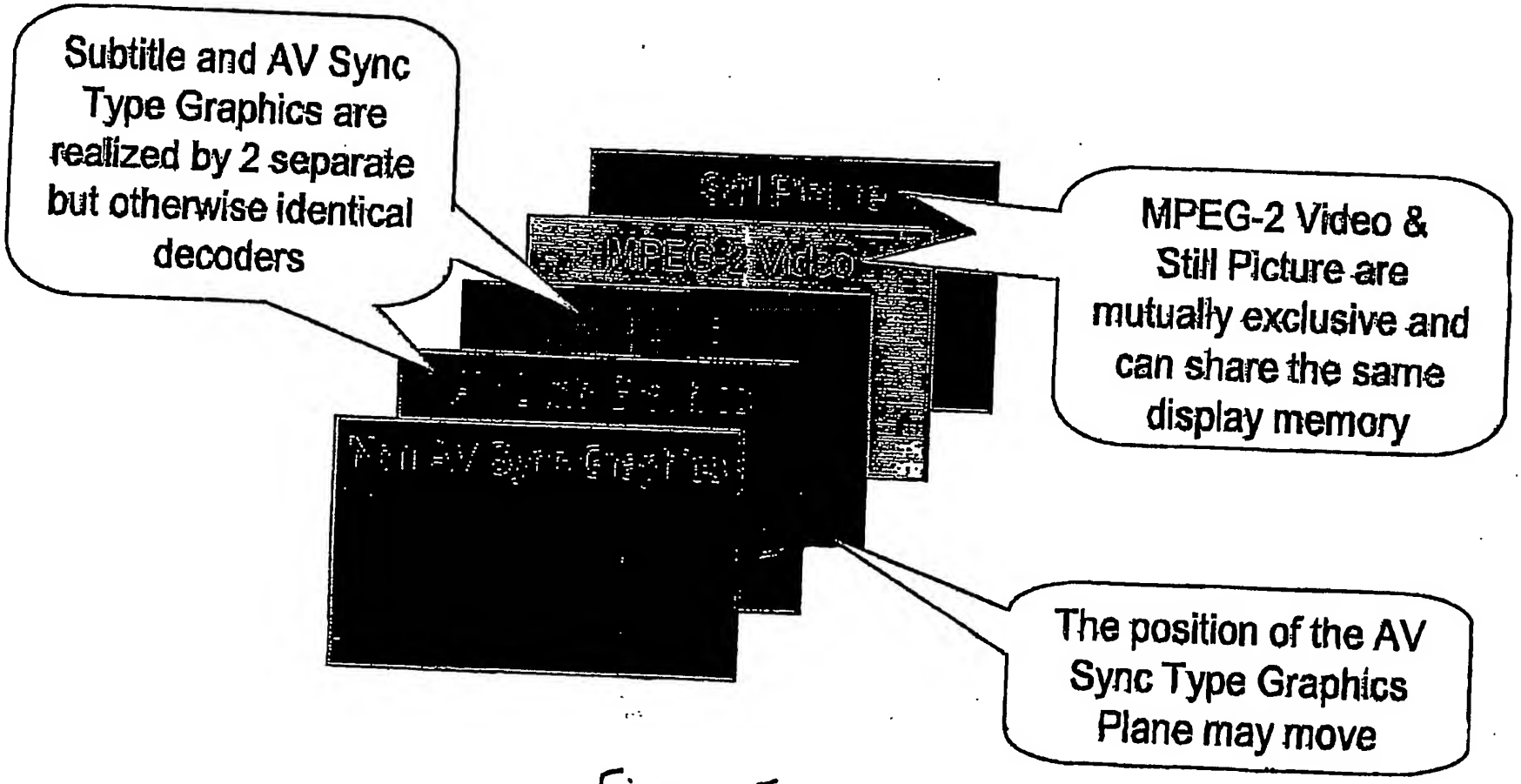


Figure 5

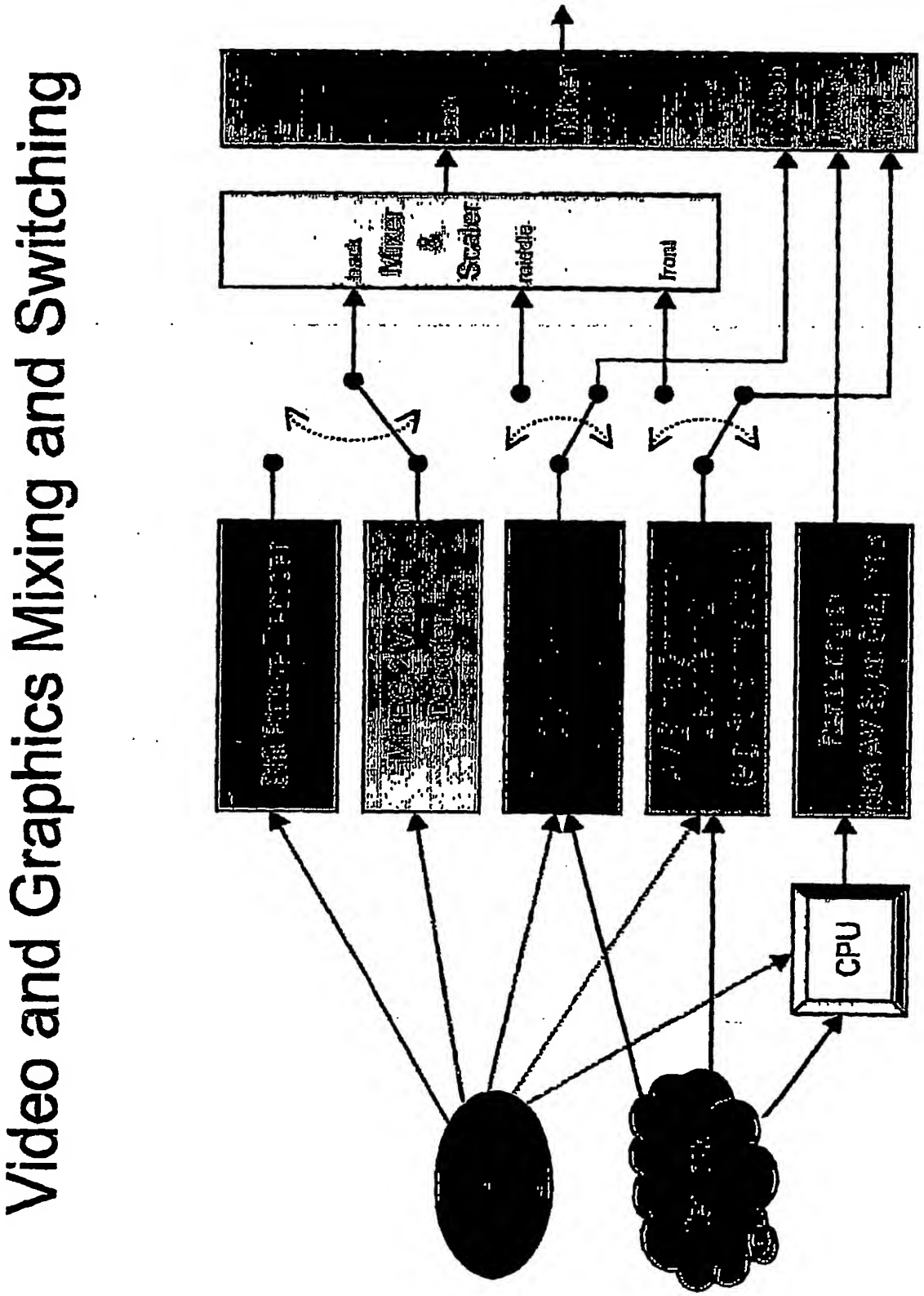

Video and Graphics Mixing and Switching
back
Mixer
Scaler
middle
front
CPU

Figure 6